Massachusetts Division of Health Care Finance and Policy
Application for All-Payer Claims Database (APCD) Data

Applications for APCD data must meet the requirements set forth in regulation **114.5 CMR 22.00: Health Care Claims Data Release** and any Administrative Bulletins promulgated under this regulation. The regulation and bulletins are available online at http://www.mass.gov/eohhs/gov/departments/hcf/regulations.html.
Information provided on pages 1-4 of this application will be posted on the internet for public comment.

A. APPLICANT INFORMATION	
Applicant Name:	Ashish K. Jha
Title:	Professor of Health Policy and Management
Organization:	Harvard School of Public Health
Project Title:	Understanding High-Cost Patients in Massachusetts
Date of Application:	February 21, 2013
Project Objectives (240 character limit)	To help policymakers understand who high-cost patients are, what types of costs they incur, and how these costs may be modified in order to control health spending in the Commonwealth of Massachusetts.
Project Research Questions	1. Identify and characterize costs and spending patterns among high-cost patients in Massachusetts using contemporary data 2. Identify the physicians, physician groups, and hospitals that care for these high-cost patients, and the communities in which these patients and providers are primarily based

B. DATA REQUESTED

1. PUBLIC USE			
File	SINGLE USE* '08 – '09 – '10	REPEATED USE* '08 – '09 – '10	MULTIPLE USE* '08 – '09 – '10
Medical Claims	X X X	☐ ☐ ☐	☐ ☐ ☐
Pharmacy Claims	X X X	☐ ☐ ☐	☐ ☐ ☐
Dental Claims	☐ ☐ ☐	☐ ☐ ☐	☐ ☐ ☐
Membership Eligibility	X X X	☐ ☐ ☐	☐ ☐ ☐
Provider	X X X	☐ ☐ ☐	☐ ☐ ☐
Product	X X X	☐ ☐ ☐	☐ ☐ ☐

2. RESTRICTED USE			
File	SINGLE USE* '08 – '09 – '10	REPEATED USE* '08 – '09 – '10	MULTIPLE USE* '08 – '09 – '10
Medical Claims	X X X	☐ ☐ ☐	☐ ☐ ☐
Pharmacy Claims	X X X	☐ ☐ ☐	☐ ☐ ☐
Dental Claims	☐ ☐ ☐	☐ ☐ ☐	☐ ☐ ☐
Membership Eligibility	X X X	☐ ☐ ☐	☐ ☐ ☐
Provider	X X X	☐ ☐ ☐	☐ ☐ ☐

Product	X X X	☐ ☐ ☐	☐ ☐ ☐

* The Division reserves the right to change proposed "use level" after review of this application.

Definitions:

- **Single Use**: Use of the data for a project or study.
- **Repeated Use**: Use of the data as an input to develop a report or product for sale to multiple clients or customers provided that it will NOT disclose APCD data. Examples include: development of a severity index tool, development of a reference tool used to inform multiple consulting engagements where no APCD data is disclosed.
- **Multiple Use**: Use of the data to develop a product or service that will be sold in the marketplace and will disclose APCD data. Examples include: a benchmark report produced by analyzing APCD data, a query tool to ease access to APCD data.

3. **Filters:** If you are requesting data elements from the Restricted Use dataset, describe any filters you are requesting to use in order to limit your request to the minimum set of records necessary to complete your project. (For example, you may only need individuals whose age is less than 21, claims for hospital services only, or only claims from small group products.)

File	Data Element(s)	Range of Values Requested
Medical Claims		
Pharmacy Claims		
Dental Claims		
Membership Eligibility		
Provider		
Product		

4. **Restricted data elements:** If you are requesting Data Elements from the Restricted Use dataset, list each restricted data element you are requesting on the attached Data Element List and explain why you need access to EACH Restricted Use data element for your project. Limit your request to the minimum data elements necessary to complete the project and be specific as to how each element relates to your proposed model/analytic plan. Add rows to this table as needed.

Restricted Data Element Name	**Restricted Data Element Description**	**Data File** (Medical, Pharmacy, Dental, Eligibility, Provider, Product)	**Justification** (reason this data element is necessary for your project)
MC016	Member zip	Medical	Will use to assign economic data (median income, etc.) from area resource file, because though we won't have any individual level socioeconomic data, it will be important to understand whether individuals of different income groups are bearing different

			burdens of utilization and spending
MC018	Admission Date	Medical	Need admission date to calculate readmission rates, which are an important clinical outcome, as well as to calculate length of stay, which is an important signal for utilization patterns
MC059	Date of service	Medical	Need dates to link claims together to create "episodes" of care for both acute and chronic conditions
MC060	Date of service	Medical	Need dates to link claims together to create "episodes" of care for both acute and chronic conditions
MC069	Discharge date	Medical	Need discharge dates to calculate and classify readmissions, which are an important clinical outcome, as well as to calculate length of stay, which is an important signal for utilization patterns
PC032	Date prescription filled	Pharmacy	Need dates to link claims together to create "episodes" of care for both acute and chronic conditions

C. PURPOSE AND INTENDED USE

1. Please describe the purpose of your project and how you will use the APCD.

Our first aim is to identify and characterize costs and spending patterns among high-cost patients in Massachusetts using contemporary data. Our second aim is to identify the physicians, physician groups, and hospitals that care for these high-cost patients, and the communities in which these patients and providers are primarily based. We will be particularly interested in understanding the predictors of spending that may be modifiable with targeted policy interventions.

We will use the Massachusetts All Payer Claims Database (APCD) for our work. The benefit of an all-payer database such as the Massachusetts APCD is that we can make comparisons across different types of payers; prior studies have shown that private and public insurance contracts may cross-subsidize each other (i.e. high Medicare rates are offset by lower private rates, and vice versa). Therefore, studying only one population may not give the most accurate results about how

costs and spending cluster.

The proposed work will allow policymakers to better understand who the high-cost patients are, what types of costs they incur, and how these costs may be modified. These data will empower clinical leaders and policy makers to craft policies – such as changes in co-payments, bundled payments, and shared savings programs (such as Accountable Care Organizations), among others – that will increase our ability to slow the growth in health care costs while protecting patients' health.

There are two types of metrics or benchmarks for assessing the progress and outcomes of this project. The first is very concrete: we expect to publish in major health policy journals and have at least two major presentations at national meetings. The second is less concrete but perhaps more important: we will share our findings early with key policymakers in Massachusetts (such as the Health Policy Council) and elsewhere to ensure that our results are available early and are accessible to those who can best use it. Our metric of success will be whether our findings are perceived to be novel, actionable, and ultimately, helpful to the key stakeholders in Massachusetts.

2. Please explain why completing your project is in the public interest.

Healthcare costs are one of the biggest challenges facing the Commonwealth of Massachusetts. In 2009 alone, Massachusetts spent more than $61 billion on healthcare, threatening other important programs such as education. Massachusetts has the highest per-capita healthcare spending in the nation, and estimates suggest healthcare will comprise more than half of Massachusetts' budget in fiscal year 2012. Furthermore, with its recent comprehensive healthcare reform, its efforts to control costs will be closely watched by policymakers across the nation. There is an urgent need to identify strategies to control healthcare spending in the Commonwealth.

One promising strategy is to focus on high-cost patients, the small number of patients that consume a disproportionate share of spending. According to the Congressional Budget Office, the top 5% of spenders in Medicare in 2001 incurred over 40% of the costs, and the top 20% of spenders over 80% of the costs. Average costs for the top 5% of patients were $63,000 per patient per year, compared with just $550 for the bottom 50%. Costs are similarly concentrated in the Medicaid program although the data here are not as robust. Less is known about the privately-insured, though costs in this group are likely similarly concentrated.

Currently, we have only a superficial understanding of high-cost patients. We know they are older and sicker, with more chronic diseases. Yet that is not adequately granular to help policymakers make smart decisions about where to target their efforts. Indeed, the CBO, in a recent report on why so many efforts to reduce healthcare spending have failed, suggested that our inadequate knowledge about high-cost patients has hindered our ability to effectively target programs. A detailed, data-driven roadmap of who these patients are and what drives their costs is critical to reining in healthcare spending.

This work has significant policy implications for healthcare in Massachusetts and nationally, and will help policymakers understand who the high-cost patients are, what types of costs they incur, and how these costs may be modified. These data will provide policymakers with a rigorous empirical foundation for designing interventions that might be effective, such as how best to structure payment reform and shared savings models. We believe that these data will improve the likelihood that efforts to reduce health care costs will be successful.

3. **Attach** a brief (1-2 pages) description of your research methodology. (This description will not be posted on the internet.)
   See attached "Research Methodology"
4. Has your project received approval from your organization's Institutional Review Board (IRB)?
   - ☐ Yes, and a copy of the approval letter is attached to this application
   - X No, the IRB will review the project on ____TBD_____
   - ☐ No, this project is not subject to IRB review
   - ☐ No, my organization does not have an IRB

D. APPLICANT QUALIFICATIONS

1. Describe your qualifications to perform the research described or accomplish the intended use.

The Principal Investigator for this proposal is Dr. Ashish K. Jha, MD MPH, Professor at the Harvard School of Public Health in the Department of Health Policy and Management. He has extensive experience studying healthcare costs, quality, and efficiency at the state and federal level, and is a nationally and internationally recognized expert on these topics. He also has extensive experience working with policymakers at many levels of government, which will help to ensure that our findings are communicated effectively with state officials who need this information.

Co-Investigators for this project are Atul Gawande, MD, and Karen Joynt, MD MPH, also from the HSPH. Dr. Gawande is also a nationally and internationally recognized expert on health care delivery innovation, and will bring a key perspective to the team. Dr. Joynt has worked with Dr. Jha on multiple recent projects in the areas of hospital costs and quality and has expertise in large database analysis. The team will also include E. John Orav, PhD, a statistician at HSPH with whom this research team has worked on numerous prior projects. Dr. Orav will provide key input on analytic design, as well as on the appropriate interpretation of results.

The team has worked together on many prior projects analyzing large databases to answer questions about variation in healthcare quality, variation in healthcare costs, and the types of hospitals that provide high-quality care at low costs, and is currently working together on a project examining high-cost Medicare patients. The core research team will be assisted by data analysts and research assistants throughout the project's duration.

2. Describe the software you plan to use to analyze the data and the experience that the applicant's team members have in using that software.

We will use STATA and SAS software to analyze the data. Our programmers and analysts have extensive experience analyzing large databases with this software and have worked on numerous projects requiring similar analytics using single payer data previously.

3. Attach résumés or curriculum vitae of the applicant/principal investigator, key contributors, and of all individuals who will have access to the data. (These attachments will not be posted on the internet.)

E. DATA LINKAGE AND FURTHER DATA ABSTRACTION

1. Does your project require linking the APCD to another dataset?
   YES X NO ☐

2. If yes, will the APCD be linked to other patient level data or with aggregate data (e.g. Census data)?
   Patient Level Data ☐ Aggregate Data X

3. If yes, please identify all linkages proposed and explain the reasons(s) that the linkage is necessary to accomplish the purpose of the project.

We will link the data with the NPPES file. This will allow us to then overlay hospital- and geographical-level data to accomplish the analyses described for Specific Aim 2 (see "Research Methodology" attached).

4. If yes, specify the specific steps you will take to prevent the identification of individual patients in the linked dataset.

After database linkage is complete, all data that could be used to identify patients (scrambled beneficiary identifier code) and data specifically identifying physicians (Unique Provider Identification Number, National Provider Identifier number, name, date of birth, medical school, etc.) will be removed and categorized (specialist training, board certification status, age in years, foreign vs. US medical school, etc.) such that no linkage back to the origin patient or physician, or any group of patients or physicians, remains. Although there will be no specific patient identifiers, after database linkage, all information potentially linkable by triangulation (e.g. admission and discharge date) will also be either removed or transformed and categorized (e.g. to length of stay). After linkage and categorization, no identifiable, or HIPAA-protected data will remain in the database, and none of the original identifiers will be stored in any format.

F. RE-RELEASE OF DATA

Applicants must obtain prior approval from the Division to publish reports that use APCD files. Applicants must provide the Division with a copy of any report at least 30 days prior to release to outside parties, including peer review and prepublication analysis by anyone other than the individuals named in this Application. The Division will review the report to ensure that the publication will not permit identification of an individual patient or permit identification of a specific payment by individual

payer. The Division may prohibit release of reports that may permit identification of individual patients or specific payment by individual payer.

1. Describe your plans to publish or otherwise disclose any APCD data elements, or any data derived or extracted from such data, in any paper, report, website, statistical tabulation, or similar document.

We expect specific deliverables to include two manuscripts in high-impact policy journals as well as dissemination of the findings at policy conferences and directly to policymakers. Our proposed project has significant implications for the Massachusetts healthcare system, and will inform ongoing policymaking on healthcare costs in the Commonwealth. We believe that given the salience of the topic, it will be easy for us to reach out to key policymakers and clinical leaders to share our findings.

2. Will the results of your analysis be publicly available to any interested party? Will you charge a fee for the reports or analysis? Please describe how an interested party will obtain your analysis and, if applicable, the amount of the fee.

The results of our analysis will be available publicly only in the form of publications in high-impact policy journals or at policy conferences.

3. Will you use the data for consulting purposes?
   YES ☐ NO X

4. Will you be selling standard report products using the data?
   YES ☐ NO X

5. Will you be selling a software product using the data?
   YES ☐ NO X

6. If you have answered "yes" to questions 3, 4 or 5, please (i) describe the types of products, services or studies; (ii) estimate the number and types of clients for which the data will be used and (iii) describe any rerelease of data by your clients.

G. USE OF AGENTS OR CONTRACTORS

Third-Party Vendors. Provide the following information for all agents and contractors who will work with the APCD data.

Company Name:	
Contact Person:	
Title:	
Address:	
Telephone Number:	
Fax Number:	

E-mail Address:	
Organization Website:	

1. Will the agent/contractor have access to the data at a location other than your location or in an off-site server and/or database?
   YES ☐ NO ☐

2. Describe the tasks and products assigned to this agent or contractor for this project.

3. Describe the qualifications of this agent or contractor to perform such tasks or deliver such products.

4. Describe your oversight and monitoring of the activity and actions of this agent or subcontractor.
